CLEANING CATEGORICAL VARIABLE (LITHOFACIES) REALIZATIONS WITH MAXIMUM A-POSTERIORI SELECTION

CLAYTON V. DEUTSCH*

Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB.,
Canada T6G 2G7

Abstract—Categorical variable images frequently present unrealistic small scale variations (noise) that are an artefact of measurement error or the geostatistical simulation method. In addition to their unpleasing appearance, these variations may have an impact on subsequent petrophysical property modeling and flow simulation. Thus, there is a need to clean such realizations without altering their desirable features (reproduction of local data and large scale spatial structure). This paper presents a straightforward algorithm and program for such cleaning. Existing algorithms such as erosion/dilation, quantile-transformation, or iterative simulated annealing-based approaches for image cleaning have significant limitations when dealing with more than two categories. The key idea behind the proposed method is to retain at each location the most probable lithofacies type based on the surrounding lithofacies types, the proximity to conditioning data, and any mismatch from the global target proportion. A number of examples are presented. FORTRAN source code for the proposed algorithm, available at the IAMG web site, is documented. © 1998 Elsevier Science Ltd. All rights reserved

Code available at http://www.iamg.org/CGEditor/index.htm

INTRODUCTION

Cell-based categorical variable simulation techniques are commonly applied to create lithofacies models prior to porosity and permeability modeling. The sequential indicator simulation (sisim) algorithm is widely used for a number of reasons: (1) local data are reproduced at their exact locations by construction, (2) there is an effective statistical control through variograms, (3) soft seismic data and large scale regional trends in lithofacies proportions can be accounted for, and (4) the results appear realistic for geological settings where there are no clear geologic facies objects. Alternative cell-based modeling techniques such as truncated Gaussian and simulated annealing approaches are also used. When the lithofacies appear to follow clear geometric patterns, such as sand-filled abandoned channels or lithified dunes, object-based lithofacies simulation algorithms are better suited.

One concern with cell-based lithofacies realizations is the presence of short scale variations/ noise. These noisy features often appear geologically unrealistic and make the realizations less credible to geologists. In some cases, the noise may have an impact on flow simulation and predicted reserves; a more justifiable reason to consider low pass filter algorithms for cleaning the realizations.

A second concern is that the lithofacies proportions of sisim realizations often depart significantly from their target (input) proportions. Typically, small proportions (5–10%) are poorly matched, with the simulated proportion systematically too high or too low. The main source of this discrepancy is the order relations correction (the estimated probabilities are corrected to be non-negative and sum to 1.0) (see Carr, 1994). Post-processing the realizations to honor target proportions is a convenient and sometimes necessary solution.

The general problem of image cleaning has been tackled by a number of workers in the area of image analysis and statistics (e.g., Schowengerdt, 1983; Geman and Geman, 1984; Gull and Skilling, 1985; Besag, 1986; Andrews and Hunt, 1989; Doyen and Guidish, 1989). One proposal to clean lithofacies realizations, closely related to some of these image analysis methods, is based on the concepts of dilation and erosion (Stoyan and others, 1987; Schnetzler, 1994). This approach is well suited to cleaning binary (only two lithofacies) images; however, there is no explicit control over the resulting proportions and no extension to more than two lithofacies types except for a sequence of nested binary cleanings.

Iterative or simulated annealing-type algorithms can be designed to clean lithofacies realizations (Doyen and Guidish, 1989; Deutsch, 1992; Murray, 1993). These methods can be very powerful; how-

*E-mail: cdeutsch@civil.ualberta.ca.

ever, there are a number of practical problems (1) they tend to be CPU intensive, (2) it is difficult to determine the appropriate values for a number of tuning parameters, and (3) often, a training image is required.

The quantile-transformation principle was extended by Xu and Journel (subsequently published (Journel and Xu, 1994; Xu, 1995) to clean lithofacies realizations and impose target proportions simultaneously. This approach (recalled in Section 2) works well when there is a natural nesting/ordering of the lithofacies categories, as would be the case when the categories are based on a continuous variable such as grain size or degree of alteration. However, as shown hereafter, unrealistic artifacts can result when dealing with more than two unordered lithofacies types.

In geostatistical notation, consider K mutually exclusive and exhaustive lithofacies types s_k, $k = 1$, ..., K and the following indicator transform:

$$i_k(\mathbf{u}) = \begin{cases} 1, & \text{if lithofacies } s_k \text{ prevails at location } \mathbf{u} \\ 0, & \text{if not.} \end{cases}$$
(1)

Mutual exclusion and exhaustivity entail that only one lithofacies may be present at any location $\mathbf{u}$ and that any lithofacies be in the list of the K types, that is,

$$\sum_{k=1}^{K} i_k(\mathbf{u}) = 1,$$
(2)

and

$$i_k(\mathbf{u}) \cdot i_{k'}(\mathbf{u}) = 0 \, \forall k \neq k'.$$
(3)

Let $\{i_k^{(0)}(\mathbf{u}), k = 1, ..., K, \mathbf{u} \in A\}$ be the set of originally simulated lithofacies values, conditional to n local indicator data values:

$$i_k^{(0)}(\mathbf{u}_\alpha) = i_k(\mathbf{u}_\alpha), \quad k = 1, ..., K, \alpha = 1, ..., n.$$
(4)

The problem is to calculate a set of "clean" lithofacies types $\{i_k^{(c)}(\mathbf{u}), k = 1, ..., K, \mathbf{u} \in A\}$, that (1) is similar to the original simulated image, (2) is less noisy than that original image, (3) more closely honors the target proportion of each lithofacies, and (4) still honors the conditioning data.

The quantile-transformation procedure will be recalled and, then, a technique based on *maximum a-posteriori selection* (MAPS) will be proposed. The key idea behind MAPS is that the lithofacies type at each location is replaced by the most probable lithofacies type based on a local neighborhood. The probability of each lithofacies type is based on (1) closeness to the location, (2) whether or not the value is a conditioning datum, and (3) mismatch from the global target proportion. In concept, the approach is similar to other, published, image processing algorithms (Schowengerdt, 1983; Besag, 1986; Doyen and Guidish, 1989) however, the

details of the weighting and the one-pass approach are unique.

A number of examples are presented to illustrate the applicability of the proposed MAPS algorithm. A FORTRAN 77 program is presented.

QUANTILE TRANSFORMATION

As explained by Journel and Xu (1994), the main goal of the quantile transformation procedure is to correct the proportions of different lithofacies classes. A by-product of that correction procedure is a cleaner, less-noisy, realization. The first step in quantile transformation is to order the K lithofacies from 1 through K. Transformation from original class k will likely be to $k - 1$ or $k + 1$, therefore, the ordering of lithofacies should mimic that encountered in the reservoir. For the hierarchy implied by the specification of s_k, $k = 1, ..., K$, define the cumulative proportion values (or cdf values) $\Pi_k^{(0)}$ from the proportions $p_k^{(0)}$ of originally simulated classes:

$$\Pi_k^{(0)} = \sum_{k'=1}^{k} p_{k'}^{(0)} \in [0, 1], \quad k = 1, ..., K$$
(5)

with $p_k^{(0)} = Prob\{I_k^{(0)} = 1\} \in [0, 1]$, $k = 1, ..., K$ and $\Sigma_k p_k^{(0)} = 1$.

Correcting the realization $\{i_k^{(0)}(\mathbf{u}), k = 1, ..., K, \mathbf{u} \in A\}$ to honor the target probabilities (p_k, $k = 1, ..., K$) requires ties to be broken. Values nearer to borders between regions of different lithofacies should be considered for re-allocation to another class. An "average" lithofacies value $b(\mathbf{u})$ will be defined at each location $\mathbf{u}$ as a moving window average (template W containing n nearest neighbors) of the category indices across all indicator values, that is:

$$b(\mathbf{u}) = \frac{1}{n(\mathbf{u})} \sum_{\mathbf{u}' \in W(\mathbf{u})} \sum_{k'=1}^{K} k' i_{k'}^{(0)}(\mathbf{u}'),$$
(6)

where $n(\mathbf{u})$ is the number of values in the template $W(\mathbf{u})$ centered at $\mathbf{u}$. The $b(\mathbf{u})$, $\mathbf{u} \in A$ values are ranked to obtain the ranks $r(\mathbf{u})$, $\mathbf{u} \in A$ from 1 to N, where N is the number of originally simulated lithofacies data. In the original presentation (Xu, 1995) the N values were partitioned into K subsets corresponding to the original indicators $i_k^{(0)}$, $k = 1, ..., K$. The ranking was then done within each subset to arrive at a global ranking. Ranking the data within subsets places more emphasis on closeness to the original simulated realization and less emphasis on smoothing; a decision best made after looking at the results for both approaches. Based on the ranking, the locations can be reclassified to obtain the correct proportions, that is,
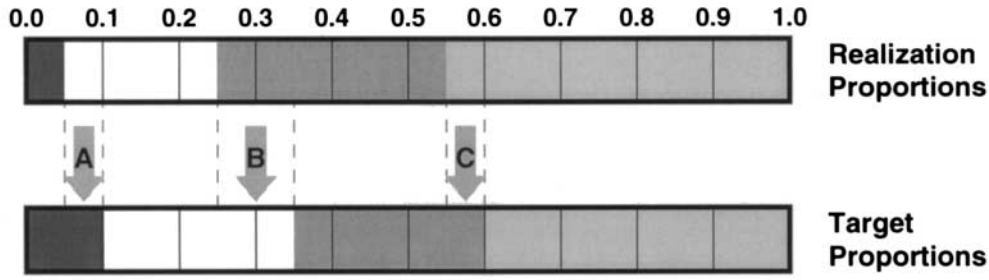
Figure 1. Schematic illustration of quantile transformation procedure to enforce reproduction of target lithofacies proportions

(1) Calculate the target cdf values:

$$\Pi_k^{(c)} = \sum_{k'=1}^{k} p_{k'} \in [0, 1], \quad k = 1, \ldots, K.$$

(2) Determine K threshold values to apply to the ranked r values:

$$r_k = \Pi_k N, \quad k = 1, \ldots, K.$$

where the 0th threshold r_0 is 0, by convention.

(3) Reclassify all locations $\mathbf{u} \in A$ according to their rank $r(\mathbf{u})$ and the thresholds,

$$i_k^{(c)}(\boldsymbol{u}) = 1, \quad i_{k'}^{(c)}(\boldsymbol{u}) = 0 \forall k' \neq k \qquad \text{if } r_{k-1} < r(\mathbf{u}) \leq r_k.$$

The resulting set of indicator values $\{i_k^{(c)}(\mathbf{u}), k = 1, \ldots, K, \mathbf{u} \in A\}$ will have the correct proportions p_k, $k = 1, \ldots, K$ and will be spatially smooth thanks to the smoothing window W.

As an example, consider the $K = 4$ lithofacies illustrated on Figure 1. The realization proportions are 0.05, 0.20, 0.30 and 0.45, respectively; shown as cumulative proportions of 0.05, 0.25, 0.55 and 1.0. The target proportions are 0.10, 0.25, 0.25 and 0.40 (cumulative proportions 0.10, 0.35, 0.60 and 1.0). The quantile transformation to the target proportions involves three changes (1) 5% of lithofa-

cies 2 is transformed to 1 (see A), (2) 10% of lithofacies 3 is transformed to 2 (see B), and (3) 5% of lithofacies 4 is transformed to 3. The local averaging procedure described above is used to break the ties of the realization lithofacies types, e.g., to establish which of the original 20% of lithofacies 2 to reclassify as type 1.

The quantile-transformation results can also be made to honor local conditioning data. This algorithm, including the ability to honor local conditioning data, has been coded in the trans program in the second edition of GSLIB. Figure 2 shows an initial 100 by 100 sisim (Deutsch and Journel, 1992) unconditional realization and the result of a quantile transformation. The target proportions in the unconditional sisim run were 0.05, 0.20, 0.30 and 0.45, respectively, for four lithofacies, 1, 2, 3 and 4. The same indicator variogram was used for all categories: 10% nugget effect, 90% spherical variogram with a range of 50 distance units in the N30E direction and 20 distance units in the N120E direction.

The proportions from the first sisim realization (left side of Fig. 2) were 0.170, 0.122, 0.210 and 0.498. This poor reproduction of the input proportions is largely due to the long range of correlation and the small size of the field (lack of
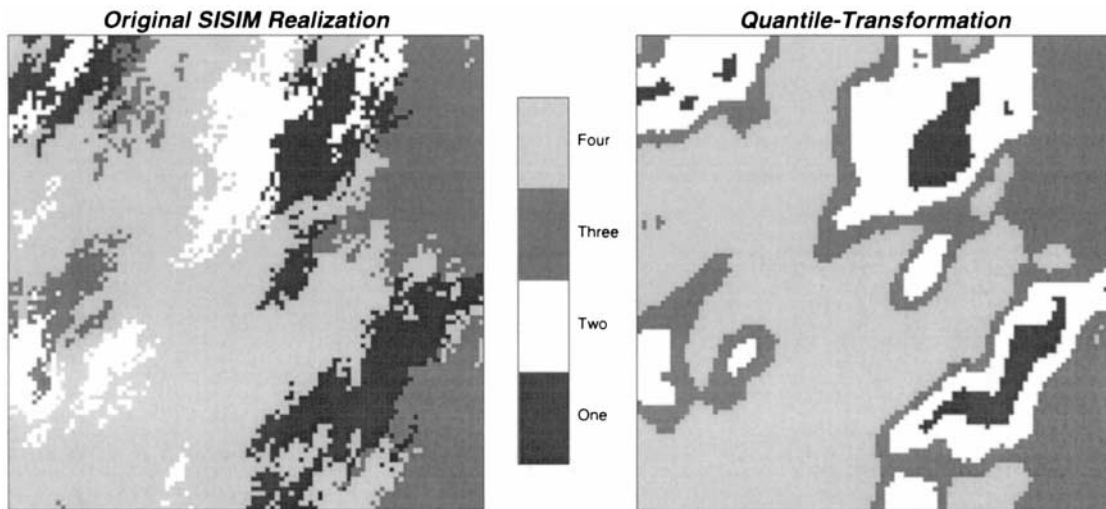

Figure 2. Sisim realization before and after quantile-transformation to match proportions

Table 1. Target `sisim` proportions for simulations with four lithofacies. Base case results are for field size A two times larger than variogram range in one direction and five times larger than range in perpendicular direction. Results for "Field 5·A" are for field size 10 times larger than range in one direction and 25 times larger in perpendicular direction

Category	Target	Realizations 1–5				
1	0.05	0.170	0.097	0.000	0.001	0.182
2	0.20	0.122	0.078	0.357	0.187	0.224
3	0.35	0.210	0.124	0.254	0.296	0.228
4	0.45	0.498	0.701	0.389	0.516	0.366
Category	Target	Field 5·A: Realizations 1–5				
1	0.05	0.080	0.056	0.052	0.059	0.057
2	0.20	0.179	0.192	0.219	0.236	0.198
3	0.30	0.323	0.261	0.274	0.269	0.330
4	0.45	0.418	0.491	0.455	0.435	0.415

ergodicity). Additional realizations were constructed where the size of the field A was increased by a factor of 5 (the field A is now 10 times larger than the largest range of correlation). The reproduction of the proportions improves but still deviates from the input proportions, see Table 1 for a summary of the results.

The program trans from the second edition of GSLIB was used to correct the proportions and to clean the `sisim` realizations. A template W of 5 pixels by 5 pixels was used in both cases (see right side of Figs 2 and 3). The cleaned realizations exactly reproduce the target proportions (by construction). The realizations are "cleaner" and show a clear "nesting" of the lithofacies categories. Unless the original lithofacies types also show nesting, this nesting is considered a limitation. More sophisticated tie-breaking procedures could be considered in the quantile-transformation algorithm; however, the inherent limitation is that lithofacies k be transformed to adjacent lithofacies types, $k + 1$ or $k - 1$. The MAPS algorithm overcomes this limitation.

MAXIMUM A-POSTERIORI SELECTION

The maximum a-posteriori selection or maps algorithm, amounts to replacing the lithofacies type at each location $\mathbf{u}$ by the most probable lithofacies type based on a local neighborhood. The probability of each lithofacies type, in the neighborhood, is based on (1) closeness to the location $\mathbf{u}$, (2) whether or not the value is a conditioning datum, and (3) mismatch from the target proportion. Figure 4 shows a schematic illustration: the cell being considered will be reclassified from code 1 to code 2 because, in the neighborhood, there are more code 2 cells thus the local probability of code 2 is higher than all other lithofacies. The lithofacies code with the maximum probability is selected; no random drawing is considered. Each node is considered independently of all others; the algorithm is not sequential. Note also that the maps algorithm does not ensure exact reproduction of the target proportions.

As before, consider an original indicator realization $\{i_k^{(0)}(\mathbf{u}), k = 1, \ldots, K, \mathbf{u} \in A\}$ at N locations
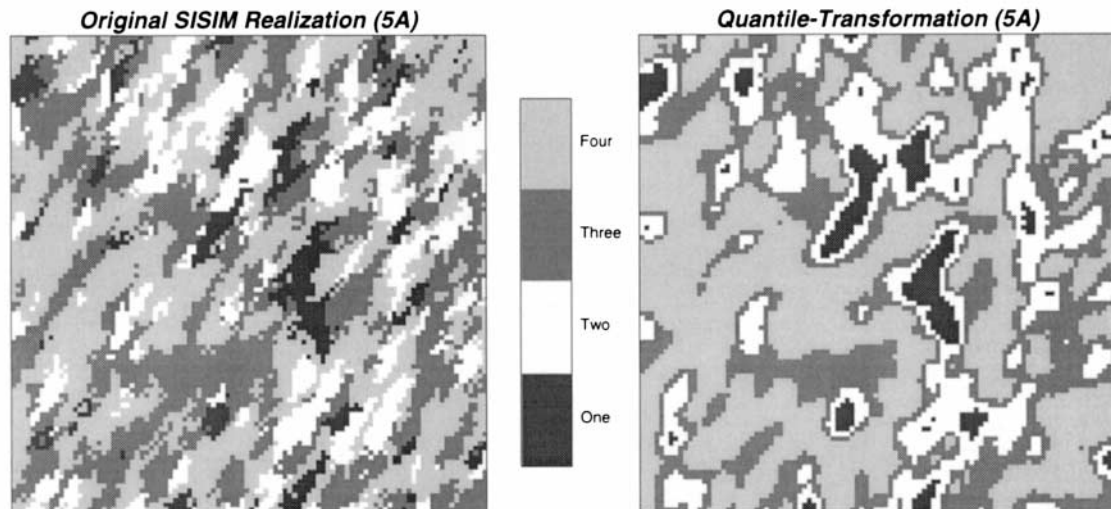

Figure 3. Sisim realization before and after quantile-transformation to match proportions. Field size is 5 times larger than that shown in Figure 2, that is, there are 100 by 100 nodes in each realization — spacing between nodes is 5 times greater in realization shown here
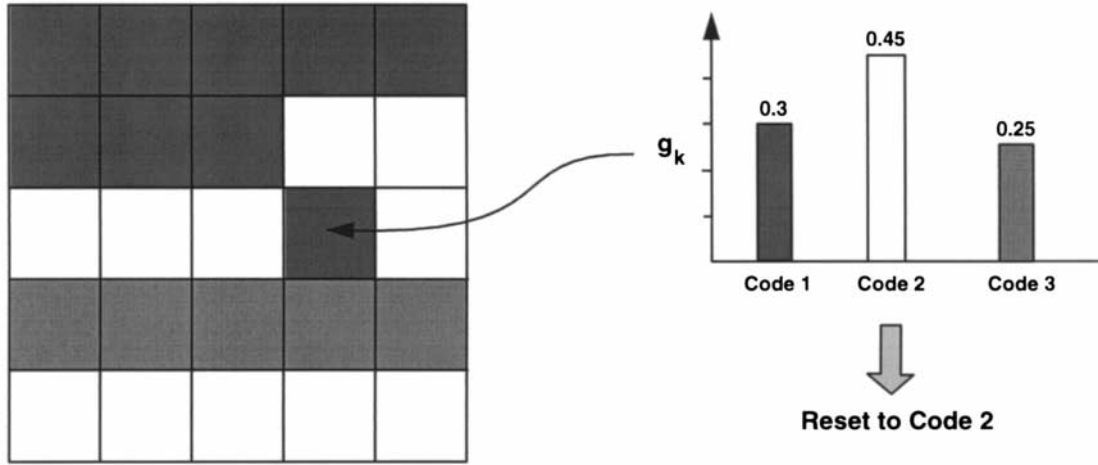
Figure 4. Schematic illustration of maps algorithm: at each location probability of each category is established on basis of nearby lithofacies codes, conditioning data, and mismatch from target proportions. Lithofacies at this location is reclassified to code 2 because, within local neighborhood, there are many code 2 lithofacies

taking one of K lithofacies types, s_k, $k = 1, \ldots, K$. There is no requirement for the lithofacies s_k, $k = 1, \ldots, K$ to be in any particular order. The proportions of each lithofacies type in the realization are $p_k^{(0)}$, $k = 1, \ldots, K$. The target proportions of each lithofacies type are p_k, $k = 1, \ldots, K$.

At each of the N locations $\mathbf{u} \in A$, calculate a local probability $q_k(\mathbf{u})$ based on a weighted combination of surrounding indicator values:

$$q_k(\mathbf{u}) = \frac{1}{S} \sum_{\mathbf{u}' \in W(\mathbf{u})} w(\mathbf{u}') \cdot c(\mathbf{u}') \cdot \frac{p_k}{p_k^{(0)}} \cdot i_k^{(0)}(\mathbf{u}') \in [0, 1],$$

$$k = 1, \ldots, K, \tag{7}$$

where S is a standardization constant to enforce $\Sigma_k q_k(\mathbf{u}) = 1.0$, $W(\mathbf{u})$ is a template of points centered at location $\mathbf{u}$, and $w(\mathbf{u}')$ and $c(\mathbf{u}')$ are weights to account for closeness to $\mathbf{u}$ and nearby conditioning data, more precisely:

● $w(\mathbf{u}')$: weight accounting for closeness to $\mathbf{u}$: to achieve realization cleaning it is necessary to establish which other lithofacies are present in the neighborhood of $\mathbf{u}$. The definition of the local neighborhood template $W(\mathbf{u})$ and the weights within it control the extent of realization cleaning. Figure 5 illustrates three possible weighting functions. The geometry and size of W and the weight function are determined by trial-and-error; see later in this paper.

● $c(\mathbf{u}')$: = weight to ensure reproduction of conditioning data $= 1.0$ at all non-data locations, and equal to C; $C \in (1, 10]$ at conditioning data
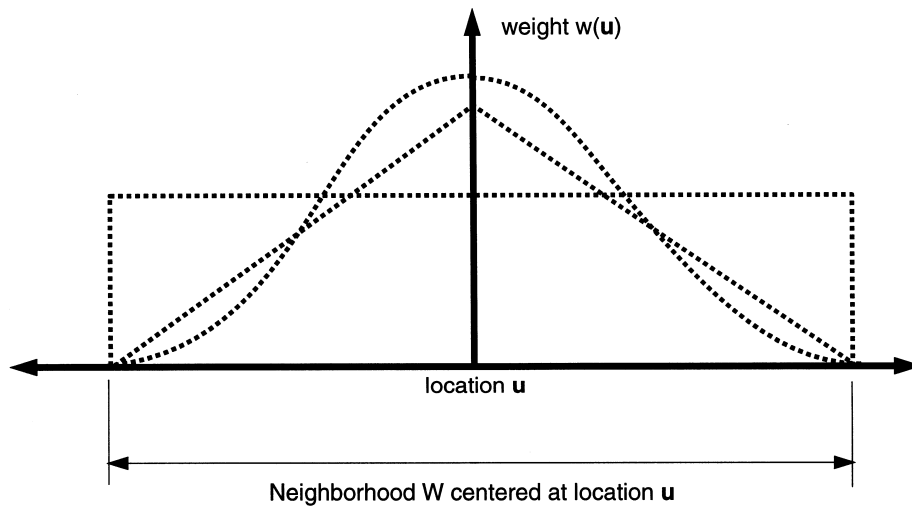

Figure 5. Three possible weight functions $w(\mathbf{u}')$ to be used within local neighborhood W for image cleaning
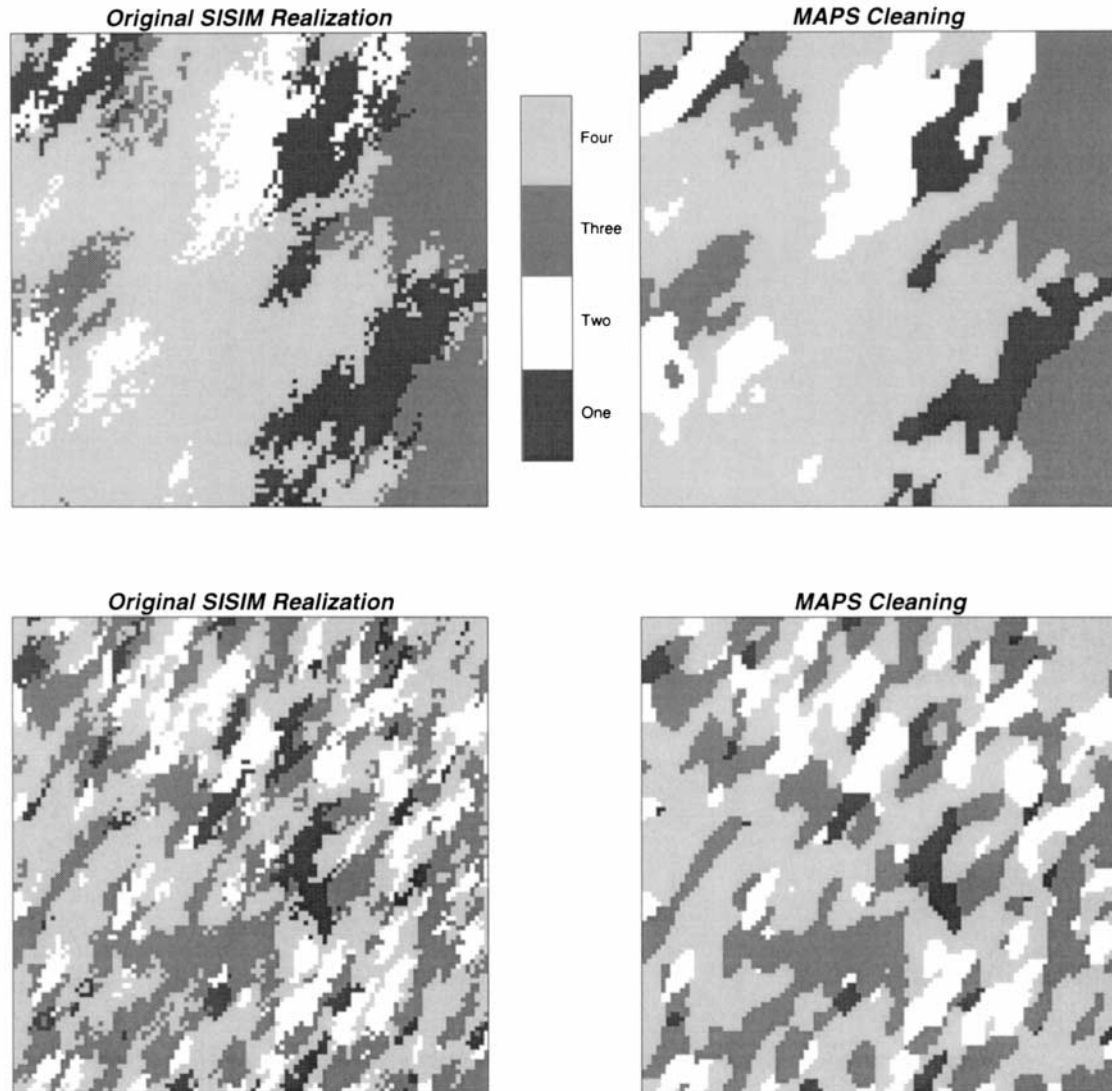
Figure 6. Two `sisim` realizations shown in Figures 2 and 3 and corresponding `maps` cleaned realizations

locations, $\mathbf{u}' = \mathbf{u}_\alpha$, $\alpha = 1, \ldots, n$. A large value of C, say $C > 10$, would entail that conditioning data have a large impact within their neighborhood, i.e., $q_k(\mathbf{u}) \approx i_k^{(c)}(\mathbf{u}')$.

The term $p_k/p_k^{(0)}$ in Equation (7) ensures that the lithofacies proportions in the cleaned image get closer to the target global proportions. The probability of a particular lithofacies q_k is decreased if the original proportion of that lithofacies is too high and increased otherwise.

Figure 6 repeats the two `sisim` realizations shown at the right of Figures 2 and 3 and shows the corresponding maps cleaned realizations. The cleaning considered a 5 by 5 pixel template W with the weights $w(\mathbf{u}')$ decreasing linearly from the central value $\mathbf{u}$. The images are cleaner and the pro-

portions of the four lithofacies types are closer to their target values, see Table 2. In the next section, we study the impact of the weighting scheme and reproduction of conditioning data.

Table 2. Target and actual proportions before and after cleaning with `maps`; proportions after cleaning are closer to target proportions

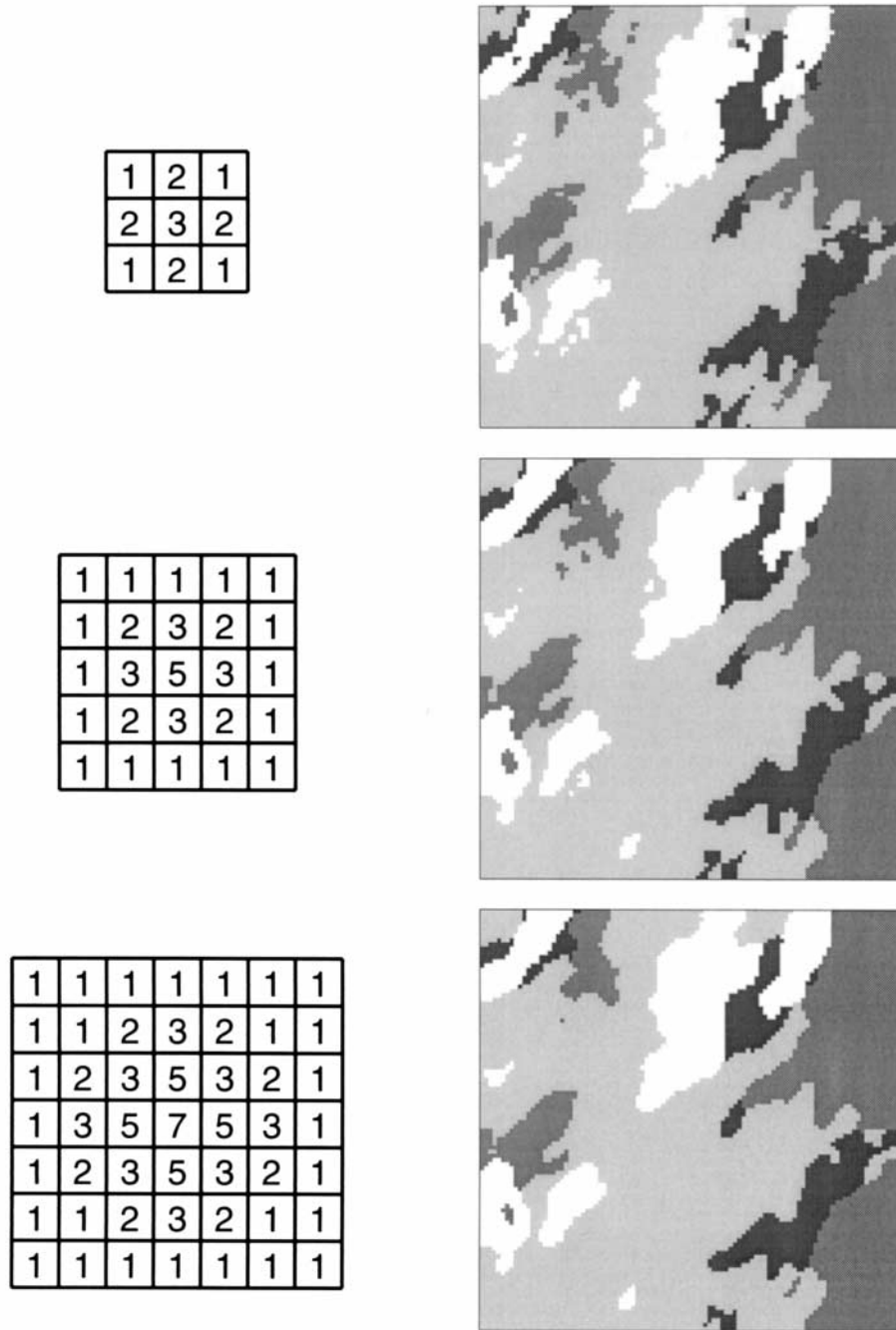| Category	Target	First realization		First 5·A realization	
		initial	final	initial	final
1	0.05	0.170	0.096	0.080	0.049
2	0.20	0.122	0.158	0.179	0.181
3	0.35	0.210	0.236	0.323	0.308
4	0.45	0.498	0.510	0.418	0.463

Figure 7. Three square templates and corresponding weights and cleaned images starting from first `sisim` realization shown in Figure 6

SENSITIVITY ANALYSIS

The size of the template W and the weighting function $w(\mathbf{u}')$ have a significant impact on the "cleanliness" of the results. In general, the image will appear smoother/cleaner when a larger smoothing window and more uniform weights are considered. Figure 7 illustrates how the first `sisim` realization shown on Figure 6 is cleaned with three different templates W and weights $w(\mathbf{u}')$. Anisotropy can be introduced to the template W and weights

$w(\mathbf{u}')$ to avoid over smoothing of original anisotropic features.

Conditioning data are enforced by the weights $c(\mathbf{u}')$ within the local neighborhood; $c(\mathbf{u}')$ is equal to 1.0 at all non-data locations, and equal to a larger value C at conditioning data locations, $\mathbf{u}' = \mathbf{u}_{\alpha}$, $\alpha = 1, \ldots, n$. Figure 8 shows two `sisim` realizations before and after cleaning. The well conditioning data are reproduced in both cases. The region on the first realization shown by the circle is
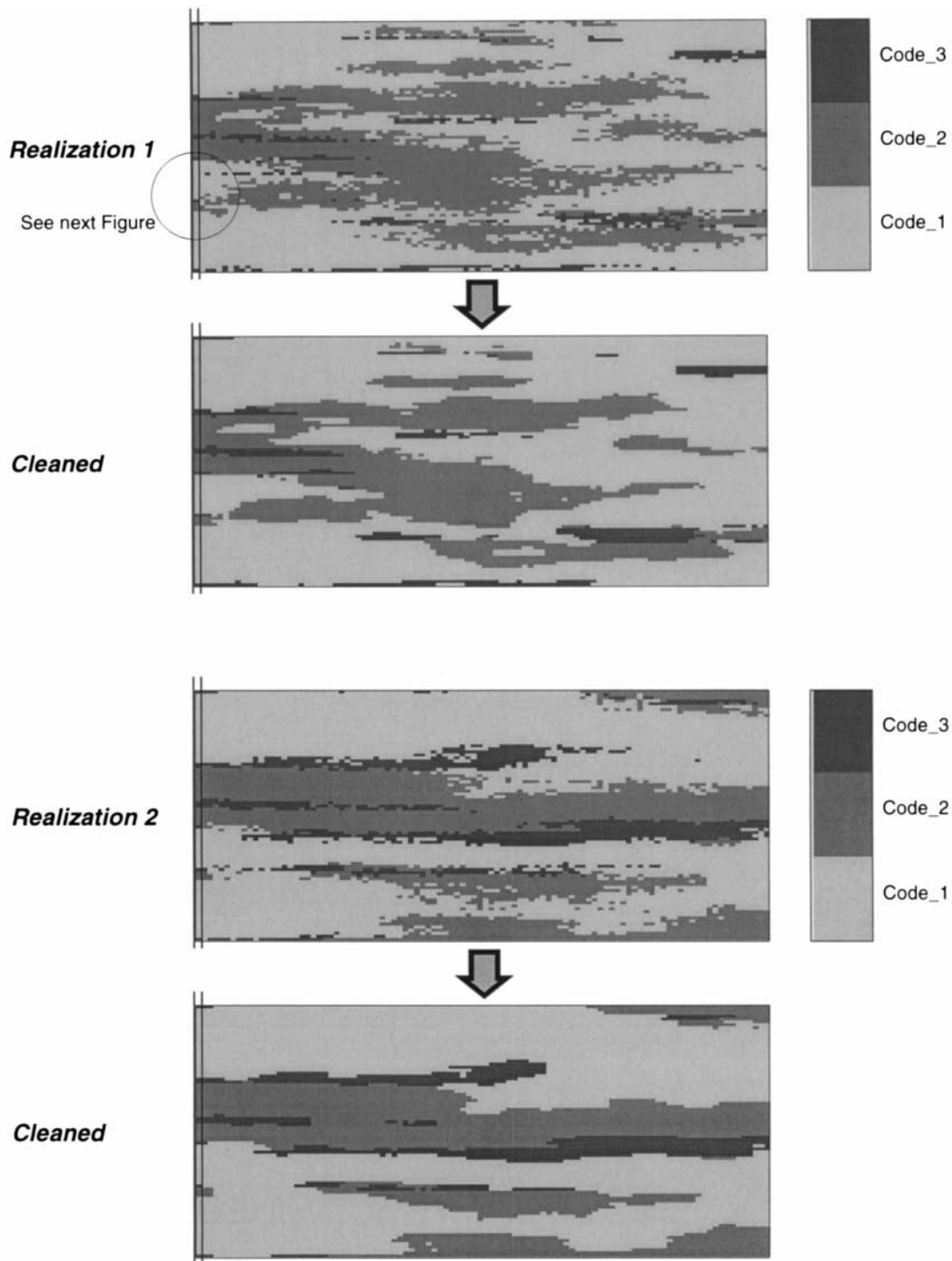
Figure 8. Two `sisim` realizations before and after cleaning. Note reproduction of conditioning data along vertical string at left

expanded and shown at the left of Figure 9; keeping all parameters but C fixed, the impact of increasing C from 2 to 5 is observed. The lithofacies observed at the vertical well are propagated continuously away from the well. The extent of the propagation depends on the magnitude of C and the size of the cleaning template W; a conditioning datum can affect only those nodes within the template W.

The term $p_k/p_k^{(0)}$ in Equation (7) leads the cleaned image toward the target proportions. Table 2 shows how the cleaned proportions are closer to the target proportions. As another example, 100 unconditional
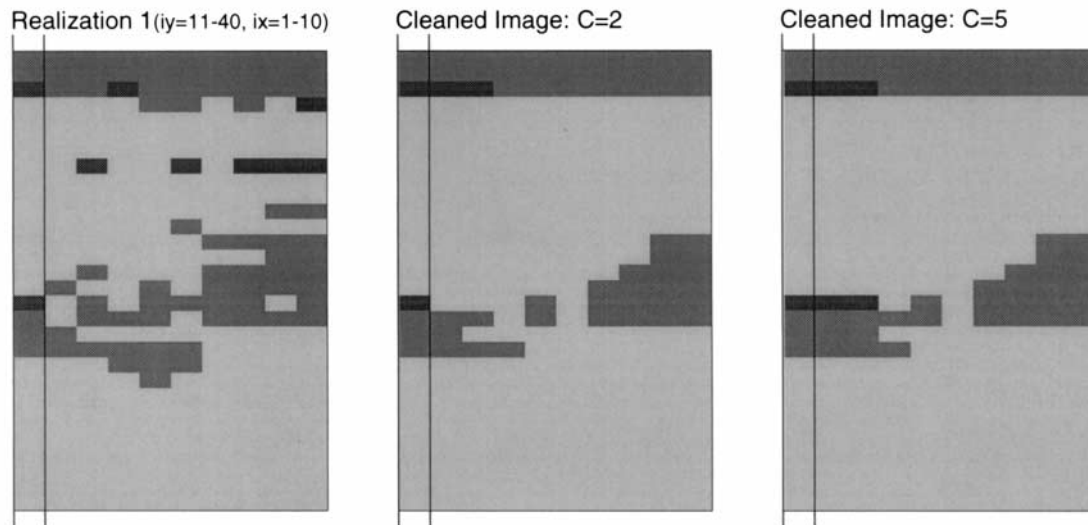
Figure 9. Portion of first `sisim` realization shown in Figure 8 and two cleaned versions. Factor *C* for conditioning is 2 in first case and 5 in second case; all other parameters are unchanged

`sisim` realizations were generated with two lithofacies types with global proportions 0.1 and 0.9, 100 by 100 cells, and isotropic correlation length of 33 cells. Figure 10 shows the proportions of the lithofacies in the 100 `sisim` realizations before and after cleaning. The target proportions are more closely reproduced. In general, variability in lithofacies proportions is an inherent aspect of uncertainty and perfect reproduction of target proportions is not a critical goal.

One could devise measures of similarity between the original and cleaned images, e.g. the proportion of cells that have changed. Quantitative measures necessarily favor one particular aspect of the problem such as (1) closeness to the original image, (2) reproduction of target proportions, or (3) smoothness or cleanliness of the resulting image. For this reason, no quantitative measures of image cleaning are presented.

As a final example, Figure 11 shows two sequential indicator simulation lithofacies models before and after cleaning.

PROGRAM PARAMETERS

The GSLIB programs have been used throughout this paper for calculations and supporting graphics. A general purpose "MAPS cleaning" program `maps` was written to clean lithofacies realizations in 2D/3D accounting for local data with arbitrary weighting terms.

The `maps` program follows GSLIB conventions. The parameters required for the program are listed below and shown in Figure 12:

● *datafl*: the input file containing the categorical lithofacies realization to be cleaned. The standard GSLIB/GeoEAS format is expected. Although there
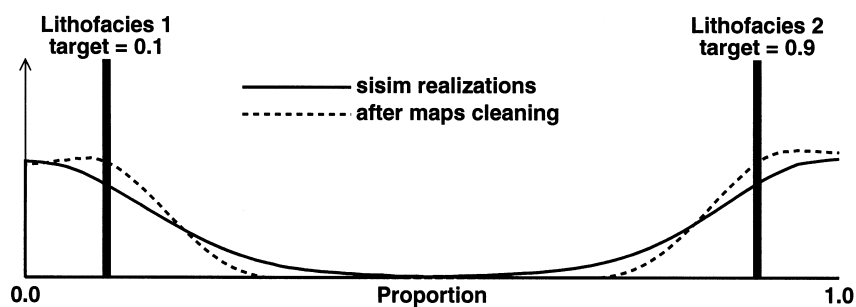

Figure 10. Proportion of lithofacies "1" and "2" in 100 `sisim` realizations before and after cleaning; target proportions were 0.1 and 0.9. Solid line is histogram of lithofacies proportions from `sisim`, that is, actual proportion of lithofacies 1 varies between 0.0 and 0.4 and actual proportion of lithofacies 2 varies between 0.6 and 1.0. Dashed line represents actual proportion after cleaning, which is closer to target proportions
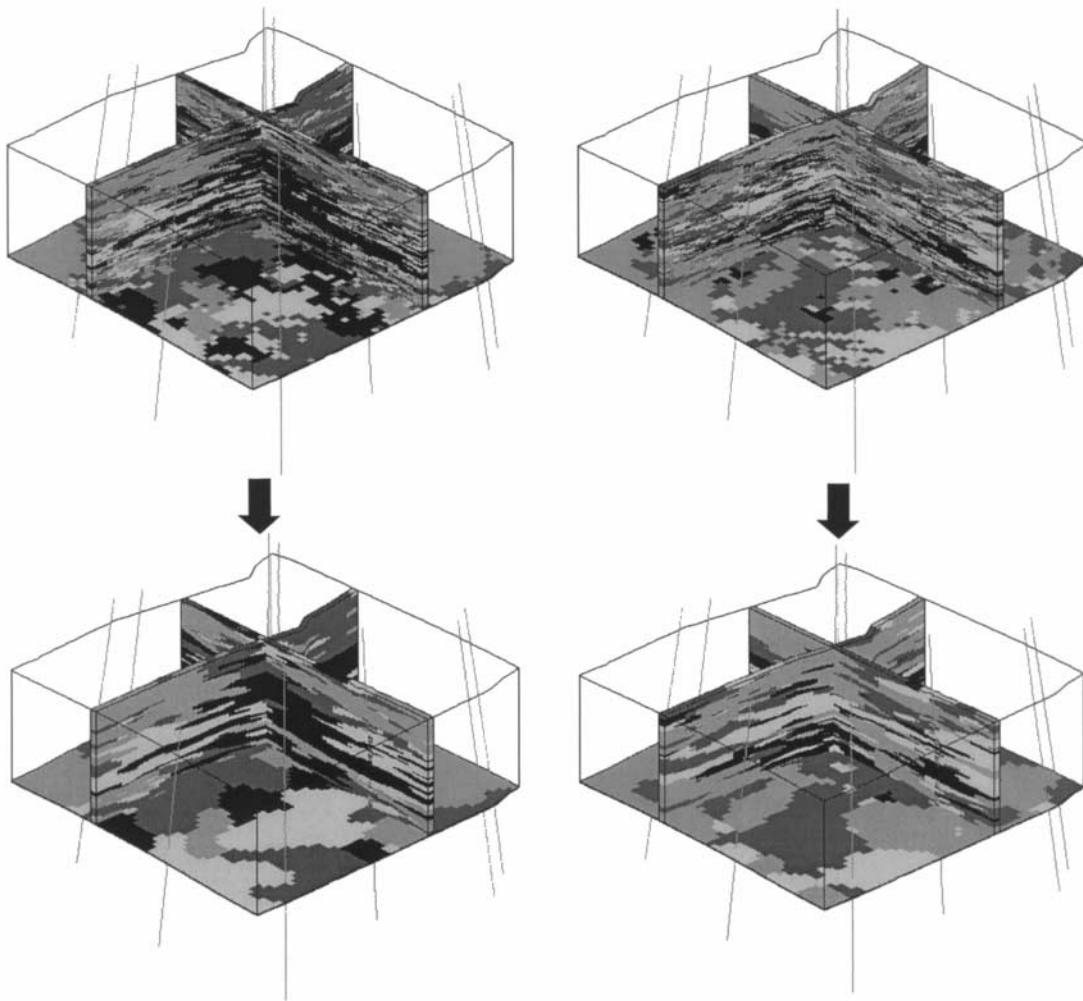
Figure 11. Two sequential indicator simulation lithofacies models before and after cleaning

```
                        Parameters for MAPS
                        ********************

START OF PARAMETERS:
sisim.out                       -file with input initial realization
maps.out                        -file for output cleaned realization
100  0.5  1.0                   -nx, xmn, xsiz
100  0.5  1.0                   -ny, ymn, ysiz
  1  0.5  1.0                   -nz, zmn, zsiz
1                               -realization number
4                               -number of categories
0    1    2    3                -    categories
0.05 0.20 0.30 0.45            -    target global proportions
1.0  1.0  1.0  1.0             -    factors for target proportions
data.dat                        -file with conditioning data
1 2 3 4                         -    columns for X, Y, Z, and category
4.0                             -    C = conditioning data weight
1                               -number of iterations
5  5  1                         -size of smoothing window W
 1.0 1.0 1.0 1.0 1.0           -    weights iy=u+2
 1.0 2.0 3.0 2.0 1.0           -    weights iy=u+1
 1.0 3.0 5.0 3.0 1.0           -    weights iy=u
 1.0 2.0 3.0 2.0 1.0           -    weights iy=u-1
 1.0 1.0 1.0 1.0 1.0           -    weights iy=u-2
```

Figure 12. Parameter file for maps program

may be multiple columns in the data file, the first column is taken as the lithofacies categories.

● *outfl*: output file for the cleaned realization. The cleaned realization will be written to a standard GSLIB file for plotting with `pixelplt...`

● *nx*, *xmn* and *xsiz*: the size of the model in the X direction.

● *ny*, *ymn* and *ysiz*: the size of the model in the Y direction.

● *nz*, *zmn* and *zsiz*: the size of the model in the Z direction.

● *isim*: the realization number to clean. `maps` only cleans one realization at a time. The program can be repeated for each realization, could be modified to automatically loop over multiple realizations, or a script file could be used to loop the program over multiple realizations.

● *ncat*: the number of categories in the realization. Any undefined categories are grouped into the first category; missing values are not allowed.

● *cat( )*: the integer coded lithofacies categories in the input data file. The values in the input data are rounded to the nearest integer.

● *pk( )*: the target proportions for each category. The local probabilities are weighted so that the cleaned result is closer to these input target proportions. The program would require modification to *not* consider target proportions (the *pk* values could, of course, be set to the realization proportions, which would entail no special weighting).

● *f( )*: the factors to weight the target proportion of each category (normally set to 1.0)

● *condfl*: file containing the conditioning data. No conditioning data will be used if this file does not exist.

● *icolx*, *icoly*, *icolz* and *icolcat*: the column numbers for the X, Y, and Z coordinate and the categorical lithofacies code in the conditioning data file.

● *C*: the factors that weights the conditioning data to ensure the cleaned realization honors the conditioning data.

● *niter*: number of iterations for the entire cleaning procedure (normally 1 is sufficient).

● *nxs*, *nys* and *nzs*: size of the W template for cleaning — these are half-window sizes, e.g., $nxs = 4$ implies that W will be 9 cells large ($2 \cdot 4 + 1$) in the X direction

● *w( )*: the weights to use within W. There are *nxs* values per row, *nys* rows for the first z slice, repeated *nzs* times for each z slice.

CONCLUSIONS

The motivation for cleaning categorical lithofacies realizations is both aesthetic and practical; noisy lithofacies realizations appear geologically unrealistic and may lead to flow simulation results with too diffusive flow character. Cleaning with classic erosion/dilation is not straightforwardly applied to realizations with more than two lithofacies types. Quantile-transformation procedures applied in cases of more than two lithofacies imparts a nesting/ordering of the lithofacies that may be unrealistic. Iterative, simulated-annealing based, cleaning algorithms often require a training image and have a large number of tuning parameters to define. The maximum a-posteriori selection `maps` procedure described here overcomes many of these problems.

In `maps`, the most probable lithofacies category is retained at each location. The corrected local probability for each category takes into account the surrounding originally simulated categories, proximity to conditioning data, and the original mismatch from the global target proportion of each lithofacies. Examples show that the approach appears flexible in its ability to clean images.

The main drawback is that there are still a number of user-defined parameters required to ensure conditioning data reproduction and to define the weighting function. A conceptual drawback of all realization cleaning algorithms is that *real* short scale variations in lithofacies may be removed for the sake of nice, clean, pretty pictures.

REFERENCES

Andrews, H. C. and Hunt, B. R. (1989) *Digital Image Restoration*. Prentice Hall, Englewood Cliffs, NJ, 238 pp.

Besag, J. (1986) On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society B* **48**(3), 259–302.

Carr, J. R. (1994) Order relation correction experiments for probability kriging. *Mathematical Geology* **26**(5), 605–621.

Deutsch, C. V. (1992) Annealing techniques applied to reservoir modeling and the integration of geological and engineering (Well Test) data. Ph.D. Dissertation, Stanford University, Stanford, California, 306 pp.

Deutsch, C. V. and Journel, A. G. (1992) *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York, 340 pp.

Doyen, P. M. and Guidish, T. M. (1989) Seismic discrimination of lithology in sand/shale reservoirs: A Bayesian approach. Expanded Abstract, SEG 59th Annual Meeting, 1989, Dallas, Texas.

Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI* **6**(6), 721–741.

Gull, S. F., Skilling, J. (1985) The entropy of an image. In: *Maximum Entropy and Bayesian Methods in Inverse Problems*, ed. C. R. Smith, W. T. Gandy, Jr., pp. 287–301. Reidel, Dordrecht.

Journel, A. G. and Xu, W. (1994) Posterior identification of histograms conditional to local data. *Mathematical Geology* **26**, 323–359.

Murray, C. (1993) Indicator simulation of petrophysical rock types. In *Geostatistics-Troia*, ed. A. Soares, Vol. 1, pp. 399–411. Kluwer, Dordrecht.

Schnetzler, E. (1994) Visualization and cleaning of pixel-based images. Master's Thesis, Stanford University, Stanford, California, 80 pp.

Schowengerdt, R. A. (1983) *Techniques for Image Processing and Classification in Remote Sensing*. Academic Press, New York, 249 pp.

Stoyan, D., Kendall, W. S. and Mecke, J., (1987). *Stochastic Geometry and its Applications*. John Wiley & Sons, New York, 345 pp.

Xu, W. (1995) Stochastic modeling of reservoir lithofacies and petrophysical properties. Ph.D. Dissertation, Stanford University, Stanford, California, 311 pp.